УДК 004.7.056.53

А. В. Роднин, В. Ю. Турчик

КОНЦЕПЦИЯ ПРИМЕНЕНИЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ В СРЕДСТВАХ ЗАЩИТЫ ИНФОРМАЦИИ БАЗ ДАННЫХ

В статье рассматриваются вопросы проектирования средств защиты информации (СЗИ) в базах данных с учетом эволюционного характера угроз информационной безопасности, обсуждаются преимущества применения интеллектуального анализа данных при создании перспективных СЗИ.

Ключевые слова: *информационная безопасность, интеллектуальный анализ данных, базы данных*

The article deals with the design of database information security systems considering evolving nature of information security threats. The benefits of the application of data mining in creating advanced information security systems are discussed.

Keywords: *information security, data mining, databases.*

Введение

Любая технология защиты информации имеет ряд ограничений, связанных со скоростью предотвращения атак на информационные системы организации. В контексте защиты БД атакой следует считать совокупность шагов, направленных на нарушение правил информационной безопасности (ИБ) путем компрометации, искажения и разглашения критически важных данных организации в обход политик ИБ.

Традиционный подход к получению актуализированного уровня защиты информации заключается в следовании требованиям законодательства (требованиям ФСТЭК, ФСБ, правилам обработки персональных данных) и стандартов (ГОСТ — государственных стандартов РФ, ISO/IEC — семейству международных стандартов серий «Информационная технология. Информационная безопасность»).

Применение технических и организационных мер обеспечивает баланс трех ключевых элементов системы управления ИБ: аппаратноого обеспечения, программного обеспечения, регламентного обеспечение.

Однако, согласно [1] в течение 2014 года 98 % организаций столкнулось с угрозами информационной безопасности. Из них 87 % организаций пострадали от утечек конфиденциальных данных, санкционированных внутренними факторами, в том числе по вине внутренних нарушителей. В связи с этим отмечается рост рынка средств класса DLP (Data Leak Prevention), которые, однако, не решают проблемы несанкционированного доступа к информации, обрабатываемой в корпоративных базах данных.

Динамика роста инцидентов ИБ говорит об их эволюционном характере. Специфика угроз такого рода заключается в сложности их формализации и, следовательно, их идентификации лицом, ответственным за ИБ. С этим связана актуальность данной работы — применение методологии интеллектуального анализа данных (ИАД) при проектировании перспективных средств защиты информации. Методы искусственного интеллекта и ИАД позволяют распознавать угрозы по совокупности неявных признаков в сочетании с набором нечетких правил и применять необходимое управляющее воздействие с целью предотвращения атаки.

Предполагается, что в результате реализации требований, перечисленных в данной работе, возможно получить необходимый уровень защищенности конфиденциальной и коммерчески значимой информации в базах данных информационных систем. Цель данной работы заключается в проектировании СЗИ, основанного на методологии ИАД применительно к действиям пользователя в базе данных.

Постановка задачи

Интеллектуальный анализ данных является одним из прогрессивных способов анализа больших объемов данных. Это процесс обнаружения и дальнейшего применения знаний или ранее неизвестной информации из уже имеющихся наборов [2].

Основными задачами ИАД являются классификация, кластеризация, ассоциация, прогнозирование, последовательность.

С точки зрения системного анализа проектируемое решение можно представить в виде следующей формальной модели:

$$<A, X, Y, K, R, S, Z>, \quad (1)$$

где A — множество состояний системы; X — набор входных объектов; Y — набор управляющих воздействий, направленных на обработку входных объектов; K — множество критериев, характеризующих поведение пользователей в системе; R — правила перехода системы из одного состояния в другое; S — правила представления данных; Z — целевое состояние системы, характеризуемое заданным уровнем защищенности.

При постановке задачи необходимо подготовить информацию о внешней среде, включающую следующие составляющие:

- отчет, содержащий данные о работе бизнес-системы;
- модель угроз информационной безопасности;
- систему критериев для выбора методов и средств, позволяющих реализовать функцию защиты информации;
- набор методов ИАД, соответствующих заданным критериям.

Пусть имеется множество ситуаций нарушения правил ИБ:

$$X = \{x_1, \ldots x_k\}, \quad (2)$$

Каждое i-е событие описывается вектором признаков:

$$x_i = \{x_i^1, \ldots x_i^n\}, \quad (3)$$

где k — количество угроз ИБ, n — количество признаков. По результатам анализа этих признаков происходит идентификация действий пользователя, а также их классификация по трем классам: «неопасные», «подозрительные», «опасные». Правила R перехода системы из состояния A_1 в состояние A_s позволяют выполнить отображение множества X на множество мер защиты информации Y.

$$Y = \{y_1,... y_m\}, \quad (4)$$

где m — количество защитных мер.

Функция классификации действий пользователя задается в следующем виде [3]:

$$F = \omega_0 + \omega_1 x_1^1 + ... + \omega_q x_k^n, \quad (5)$$

где ω_1, ω_2,... ω_q — веса независимых переменных, в поиске которых и состоит задача нахождения классификационной функции.

Далее обсуждаются требования к СЗИ, основанным на интеллектуальном анализе действий пользователя в базах данных с учетом эволюционных характеристик угроз ИБ.

Требования

1. Высокая степень интеграции с бизнес-системой заказчика

Высокий уровень системной интеграции достигается при учете необходимости проектирования межсистемных интерфейсов СЗИ — бизнес-система. Причина необходимости повышения уровня системной интеграции вызвана гетерогенностью бизнес-системы заказчика.

В данной работе предлагается рассматривать проектируемое решение как API (application programmer interface), с возможностью встраивания в системы управления базами данных от различных разработчиков. Прототипом такого СЗИ является продукт КриптоПро, разработанный на основе «OpenCrypto API». Такой подход позволил разработчикам добиться возможности встраивания данного средства криптозащиты как конструктора в различные приложения: электронную почту, браузеры, CRM-, ERP-систему и т. д. Для разных ситуаций, в которых необходимо шифрование данных, предусматривается подключение разных криптопровайдеров в виде наборов библиотек, реализующих определенные функции криптозащиты.

2. Память и возможность прогнозирования возникновения угрозы

Традиционный подход к проектированию СЗИ предполагает обеспечение следующих свойств информации в базах данных: конфиденциальность, целостность, доступность, неотказуемость, аутентичность, верифицируемость, возможность восстановления после сбоя, подотчетность, надежность.

Методология ИАД позволяет более эффективно выполнять оценку состояния наблюдаемых процессов, выявлять и ранжировать причины значимых изменений, анализировать развитие процессов и вырабатывать рекомендации по подготовке

возможных вариантов решений с прогнозом их последствий. Важной особенностью такого подхода является возможность реализовать в СЗИ эволюционные свойства адаптации, самоорганизации, обучения, возможности наследования и представления опыта экспертов ИБ в виде доступных для анализа системы нечетких правил [4].

3. Адаптивность по отношению к внешней среде

В общесистемном плане адаптация — это способность системы обнаружить целенаправленное приспосабливающееся поведение в сложившихся средах, а также сам процесс такого приспособления [5].

В рамках данной работы предлагается рассматривать угрозы нарушения конфиденциальности, целостности и доступности данных, как воздействия, обладающие эволюционным характером. В перспективных средствах защиты информации для реализации функции адаптивности применяются нейронные сети.

Нейронная сеть, обученная с помощью массива первичных данных, содержащего информацию об одной из логических функций, позволяет выявлять скрытую закономерность — вид соответствующей логической функции [6].

4. Управление событиями безопасности и формирование реакции

В теории СЗИ различают одновременное, опережающее и запаздывающее противодействие. Запаздывающее противодействие — такое противодействие, при котором реакция системы защиты начинается к моменту завершения атаки или после нее. Одновременное противодействие — то, что начинается с появлением угрозы. И, наконец, противодействие, носящее опережающий характер, означает, что реакция системы защиты начинается до начала реализации угрозы.

Управление событиями предлагается рассматривать на основе правил. Правила представляются в виде «если - то» и также используются для прогнозирования. На основе частоты встречаемости логических закономерностей делается вывод о возможном событии. Например, цепочка MD — COPYA3 — ARH — WWW — DEL ассоциируется с копированием информации из конфиденциального источника и передачей ее по каналам связи сети Интернет [7].

5. Расширенный мониторинг событий безопасности и их протоколирование

При работе с очень важными данными или при выполнении ответственных операций возникает необходимость организации контрольного журнала (audit trail), в который вносится информация обо всех событиях, происходящих в системе [8].

Типичная запись в файле контрольного журнала может содержать такую информацию:

- сам запрос (исходный текст запроса);
- номер терминала, с которого была затребована операция;

– имя пользователя, затребовавшего операцию;
– переменные отношения, кортежи и атрибуты, вовлеченные в процесс выполнения операции;
– исходные значения изменяемых данных (старые значения);
– модифицированные значения данных (новые значения).

Структурная модель решения

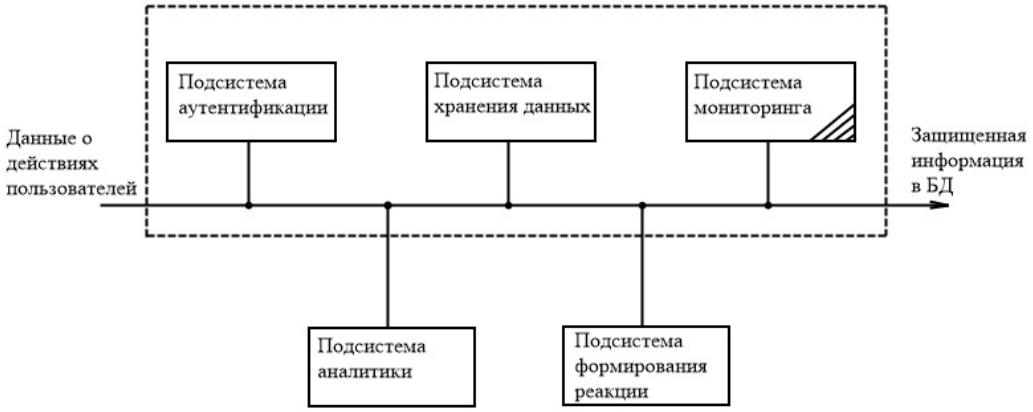

Рис. 1. Системно-структурная модель СЗИ:
В рамке представлен компилятивный прототип.
Штриховка означает модификацию традиционной системы

1. Подсистема аутентификации

Применение открытых каналов передачи данных создает потенциальные возможности для действий злоумышленников (нарушителей). Поэтому одной из важных задач обеспечения информационной безопасности при взаимодействии пользователей является использование методов и средств, позволяющих одной (проверяющей) стороне убедиться в подлинности другой (проверяемой) стороны. Обычно для решения данной проблемы применяются специальные приемы, дающие возможность проверить подлинность проверяемой стороны [9].

2. Подсистема хранения данных

Природа операций ввода-вывода, кэши баз данных, соотношение операций чтения и записи и прочие детали выглядят совершенно по-разному в системах OLTP и DSS. Кроме того, ожидания в отношении показателей времени отклика между этими системами тоже отличаются. Следовательно, схема хранения, прекрасно подходящая для базы данных одного типа, может совершенно не подходить для базы данных другого типа, поэтому для выбора правильных вариантов в этой чрезвычайно важной области, на стадии физического проектирования необходимо узнать об операционных потребностях своего приложения как можно больше [10].

3. Подсистема мониторинга и журналирования

Регистрация действий пользователя осуществляется в журналах аудита СУБД. Подсистема мониторинга и журналирования служит источником информационного сырья, которое необходимо представить в пригодной для анализа форме. Записи в журнале представляют собой некоторое количество кортежей, пригодных для анализа при условии их очистки, удаления недопустимых и ошибочных комбинаций.

Улучшить данную подсистему предлагается за счет введения более чувствительных метрик, направленных на выявление подозрительной деятельности пользователя в базе данных.

4. Подсистема аналитики

Введение данной подсистемы в СЗИ обеспечивает развитие бизнес-логики имеющейся системы. Для исследования поведения различных выборок кортежей с целью определения функции классификации необходимо использовать аппарат теории нечетких множеств и методы интеллектуального анализа данных.

5. Подсистема формирования реакции

Данная подсистема предназначена для выработки оптимального управляющего воздействия, направленного на предотвращение нарушения конфиденциальности, целостности и доступности информации, хранящейся в базе данных. При проектировании данной подсистемы применяется аппарат теории автоматов и лучшие практики (code of practice) информационной безопасности.

Выводы

К основным видам угроз ИБ относятся:

– связанные со злонамеренной модификацией параметров функционирования системы внутренним нарушителем;

– несанкционированный доступ к конфиденциальной информации, имеющейся в системе с целью ознакомления, модификации, блокирования;

– связанные с разграничением прав доступа;

– связанные с передачей информации по каналам связи и работе в сети Интернет.

Рост числа инцидентов ИБ говорит об эволюционном характере угроз. С целью определения потенциальной или реализуемой атаки предлагается использовать методики интеллектуального анализа данных. ИАД — это процесс обнаружения ранее неизвестных скрытых зависимостей и дальнейшего применения знаний из имеющихся наборов данных.

Литература

1. Лаборатория Касперского. Информационная безопасность бизнеса. Исследование текущих тенденций в области информационной безопасности бизнеса. 2014. 19 с.
2. *Han J.* Data Mining: Concepts and Techniques. Morgan Kaufmann, 2000.

3. *Барсегян А. А.* Методы и модели анализа данных: OLAP и Data Mining. СПб.: БХВ-Петербург, 2004. 336 с.
4. *Маслова Н. А.* О применении интеллектуального анализа данных для защиты информации корпоративных систем // Искусственный интеллект. 2009. № 4. С. 66–74.
5. Математика и кибернетика в экономике: словарь-справочник. М.: Экономика, 1975.
6. *Жуков В. Г.* Модель синтеза коллективов интеллектуальных информационных технологий решения задачи обнаружения инцидентов информационной безопасности // Программные продукты и системы: международный научно-практический журнал. 2014. № 1(105). С. 29–35.
7. *Маслова Н. А.* О применении интеллектуального анализа данных для защиты информации корпоративных систем. С. 68.
8. *Дейт К. Дж.* Введение в системы баз данных / 8-е изд. М.: ИД «Вильямс», 2005. 1309 с.
9. *Шаньгин В. Ф.* Информационная безопасность компьютерных систем и сетей: учеб. пособие. М.: ИД «ФОРУМ»: ИНФРА-М, 2008. 416 с.
10. *Алапати, Сэм Р.* Oracle Database 11g: руководство администратора баз данных. М.: ООО «И. Д. Вильямс», 2010. 1440 с.